Computer Vision

SS 2011

Skript

(Work in Progress)

Simon Hawe & Martin Kleinsteuber

Skript: Manuel Wolf

Inhaltsverzeichnis

1 Einführung

1.1 Was ist ein Bild?

Eine Funktion $I : \Omega \subset \mathbb{R}^2 \mapsto \mathbb{R}$ heißt monochromatisches Bild.
oft: Ω rechteckig (CCD-Sensor, Charge-coupled-device)
Bei digitalen Bildern ist sowohl Ω als auch der Wertebereich diskret und endlich,
z.B. $\Omega = [1640] \times [1480] \cap \mathbb{Z}^2, I(x, y) \in [0, 255] \cap \mathbb{Z}$

Physikalische Interpretation: $I(x, y)$ ist Bestrahlungsstärke (in $\frac{W}{m^2}$).

- Bild als Graph von I
- Digitale Bilder als Matrix
- Darstellung mit Grauwerten $\rightarrow$ („reales“ Bild)

1.2 Wie entsteht ein Bild?

1.2.1 Bild durch (dünne) Linse

1. Strahlen durch 0 werden nicht abgelenkt
2. Strahlen parallel zur optischen Achse laufen durch den Brennpunkt

$\frac{Z}{z} = \frac{B}{b} = \frac{f}{z-f}$ (Strahlensatz)
$\Rightarrow \frac{z}{Z} = \frac{z}{f} - 1$
$\Rightarrow \frac{1}{Z} + \frac{1}{z} = \frac{1}{f}$ (Gleichung für die dünne Linse)

1.2.2 Einfaches Kameramodell: Die Lochkamera

Annahme: Öffnung $\approx 0 \longrightarrow$ alle Strahlen gehen durch das optische Zentrum, d.h. Bildpunkte sind immer in der Brennebene.

$p \in \mathbb{R}^3$ mit Koordinaten $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ relativ zu 0, $p' = \begin{bmatrix} x' \\ y' \\ f \end{bmatrix}$ auf Brennebene.

$\frac{y'}{f} = \frac{y}{z} \Rightarrow y' = \frac{f}{z} \cdot y$
Analog: $x' = \frac{f}{z} \cdot x$

1.2.3 Perspektive Projektion

Durch Identifikation der Brennebene mit $\mathbb{R}^2$ hat das Bild von p die Koordinaten $\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}$

Definition 1.1. Die Abbildung $\Pi : \mathbb{R}^3 \to \mathbb{R}^2$, $\begin{bmatrix} x \\ y \\ z \end{bmatrix} \mapsto \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}$ heißt *ideale perspektive Projektion.*

Diskussion: Wann ist diese Gleichung gültig?
$\to$ Bei zu kleinem z gehen die Koordinaten der Bildpunkte gegen unendlich; das Modell ergibt keinen Sinn mehr.

1.3 Homogene Koordinaten

Definition 1.2. Sei $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$. Dann heißen $\begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1}$ die *homogenen Koordinaten* von x.

Frage: Was sind die Urbilder von $\begin{bmatrix} x \\ y \end{bmatrix}$ unter der idealen perspektiven Projektion?

$$\Pi^{-1} \begin{bmatrix} x \\ y \end{bmatrix} = \left\{ \lambda \cdot \begin{bmatrix} x \\ y \\ f \end{bmatrix} | \lambda \in \mathbb{R} \right\}$$

Jedem Punkt im $\mathbb{R}^n$ kann eine Gerade (homogene Koordinaten) im $\mathbb{R}^{n+1}$ zugeordnet werden. Umgekehrt klappt es auch, falls nicht für alle Punkte der Geraden gilt $x_{n+1} = 0$.

$$\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} I_n | 0 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{bmatrix}, I_n : (n \times n)\text{-Einheitsmatrix}$$

$=: \Pi_0$ heißt „kanonische Projektion".

Beispiel 1.3. Die perspektive Projektion eines Punktes liefert $\begin{bmatrix} x' \\ y' \end{bmatrix} = \frac{f}{z} \begin{bmatrix} x \\ y \end{bmatrix}$
bzw.: $z \cdot \begin{bmatrix} x' \\ y' \end{bmatrix} = f \cdot \begin{bmatrix} x \\ y \end{bmatrix}$.

Der Übergang zu homogenen Koordinaten liefert:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{z}x \\ \frac{f}{z}y \\ 1 \end{bmatrix} = \frac{1}{z} \begin{bmatrix} fx \\ fy \\ z \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{z} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Pi_0 \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Wir erhalten somit das *geometrische Modell einer idealen Kamera*:

$$z \cdot \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Pi_0 \cdot \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

2 Feature Point Extraction

Idee:
Zwei Bilder von einem Objekt. Finde Korrespondenzen eines markanten Merkmals (Feature) des ersten Bildes im zweiten.
Mögliche Anwendungen: Tiefeninformationen aus Stereokameras (analog zum menschlichen Auge), Bewegung zwischen zwei Bildern, ...
Wie findet man solche Merkmale? $\rightarrow$ Kanten und Ecken.

2.1 Corner and Line Detection

Idee:
kleiner Bildausschnitt, Verschiebung um u.
Inmitten Fläche: Bild ändert sich nie.
Am Rand: Bild ändert sich, außer Bild wird entlang der Kante verschoben.
An Ecke: Bild ändert sich immer.

Benötigte Mathematik:

- LinAlg:
  - Eigenwerte
  - Eigenwertzerlegung reeller symmetrischer Matrizen
  - Positiv-semidefinite Matrizen
- Calculus
  - Taylor-Approximation 1. Ordnung

2.1.1 Wiederholung Lineare Algebra

Lineare Abbildung $A : \mathbb{R}^n \rightarrow \mathbb{R}^m A \in \mathbb{R}^{m \times n}, x \mapsto A \cdot x$

$Ker\ A = \{x \in \mathbb{R}^n \mid Ax = 0\}$
$Bild\ A = \{Ax \mid x \in \mathbb{R}^n\}$

Dimensionsformel: $\boxed{dim\ Ker\ A + \underbrace{dim\ Bild\ A}_{= \text{Rang von A}} = n}$

Definition 2.1. Transponierte von A: $A^T \in \mathbb{R}^{n\times m}$
Eigenschaften: $(Bild\ A)^\perp = Ker\ A^T$ ($\perp$: orthogonales Komplement)

Betrachte $A \in \mathbb{R}^{n\times n}$.
$A : \mathbb{R}^n \to \mathbb{R}^n$

Definition 2.2. Falls $A \cdot x = \lambda \cdot x$ für ein $\lambda \in \mathbb{R}, x \in \mathbb{R}^n \setminus \{0\}$,
dann heißt x Eigenvektor von A zum Eigenwert λ.

Beobachtung: Nicht jede *reelle* Matrix hat *reelle* Eigenwerte und -vektoren.
(z.B. $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$)

Definition 2.3. Falls eine Basis des $\mathbb{R}^n$ aus Eigenvektoren besteht, z.B. $[v_1, ..., v_n] =: V$:

$$A \cdot [v_1 \dots v_n] = [A \cdot v_1 \dots A \cdot v_n] = [\lambda_1 \cdot v_1 \dots \lambda_n \cdot v_n] = [v_1 \dots v_n] \cdot \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{=:D}$$

$\Rightarrow A = V \cdot D \cdot V^{-1}$ heißt Eigenwertzerlegung von A.

Definition 2.4. $A \in R^{n\times n}, A = A^T$ (symmetrisch)
und $x^T Ax \geq 0, \forall x \in \mathbb{R}^n \to$ A heißt positiv semidefinit.
Falls $x^T Ax > 0, x \neq 0 \to$ *positiv definit*

Beispiel 2.5. $\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ ist positiv definit. Zur Überprüfung: Berechne Eigenvektoren.

Satz: Wenn A positiv definit, dann gilt:
$\mathbb{R}^n$ besitzt Basis aus Eigenvektoren von A.
Alle Eigenwerte von A sind > 0 (semidefinit: ≥ 0).
$A = A^T$

Allgemein: $A = A^T$, dann besitzt A eine reelle Eigenwertzerlegung.
Eigenvektoren zu unterschiedlichen Eigenwerten sind orthogonal, d.h. V kann so gewählt werden, dass: $A = V \cdot D \cdot V^{-1}, v_i^T \cdot v_j = 0$
oBdA. $\|v_i\| = 1$, d.h. V ist orthogonale Matrix ($V^T V = I_n$). $\Rightarrow V^{-1} = V^T$
$\hookrightarrow A = V \cdot D \cdot V^T$, V orthogonal, $V \in O(n) = \{V \in \mathbb{R}^{n\times n} | V^T V = I_n\}$

2.1.2 Harris-Edge/Corner-Detector

Sei $I : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$ ein Bild. Sei $W(x) \subset \Omega$ ein „Fenster“ um $x \in \Omega$.
Betrachte: $S(u) = \sum_{\tilde{x}\in W(x)} (I(\tilde{x}) - I(\tilde{x} + u))^2$
$I(\tilde{x} + u) \overset{Taylor}{\approx} \nabla I(\tilde{x}) \cdot u + I(\tilde{x})$, $\nabla I(\tilde{x})$: Gradient von I bei $\tilde{x}$.

$S(u) \approx \sum_{\tilde{x} \in W(x)} (\underbrace{\nabla I(\tilde{x})}_{\left[\frac{\partial I}{\partial x_1}(\tilde{x}) \quad \frac{\partial I}{\partial x_2}(\tilde{x})\right]} \cdot \underbrace{u}_{\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}})^2$

$= \sum_{\tilde{x} \in W(x)} u^T \cdot \nabla I(\tilde{x})^T \cdot \nabla I(\tilde{x}) \cdot u$

$= \underbrace{u^T \cdot \left(\sum_{\tilde{x} \in W(x)} \nabla I(\tilde{x})^T \cdot \nabla I(\tilde{x}) \right)}_{=:G} \cdot u \quad , G(x) \in \mathbb{R}^{2 \times 2}$ ist symmetrisch und *pos. semidefinit*.

Beobachtung:

homogene Fläche: $G(x) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

Kante: EW von $G(x) : \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix}$

$G(x) = V \cdot \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \cdot V^T \hookrightarrow S(u) \approx \underbrace{u^T \cdot V}_{=:\tilde{u}} \cdot \begin{bmatrix} \lambda & 0 \\ 0 & 0 \end{bmatrix} \cdot \underbrace{V^T \cdot u}_{=:\tilde{u}^T}$

$= \tilde{u}_1^2 \cdot \lambda_1 + \tilde{u}_2^2 \cdot \lambda_2$

Falls $\lambda_1, \lambda_2 = 0$: $S(u) \approx 0$ für alle $\tilde{u}$, somit für alle u.
$\longrightarrow$ Homogene Fläche.

Falls $\lambda_1 \neq 0, \lambda_2 = 0$: $S(u) \approx 0$ für $\tilde{u} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \cdot s, s \in \mathbb{R}$, g.d.w. u Eigenvektor von $G(x)$ zum Eigenwert 0 ist.
$\longrightarrow$ Kante entlang $x + u$.

Falls $\lambda_1 \neq 0, \lambda_2 \neq 0$: $S(u) \approx 0$ nur für $u = 0$
$\longrightarrow$ Ecke.

Anmerkung: Es kann hier zusätzlich ein Gewichtungsfaktor $w(\tilde{x})$ verwendet werden.

(sehr naive) Approximierung des Gradienten: $f'(x) = h \overset{lim}{\rightarrow} 0 \ \frac{f(x+h)-f(x)}{h}$
$\overset{h=1}{\hookrightarrow} \frac{\partial}{\partial x} I(x,y) \approx I(x+1,y) - I(x,y)$

3 Bild, Urbild und Co-Bild von Linien und Punkten

Definition 3.1. $x, y \in \mathbb{R}^n$ 0 schreibe $x \sim y$ falls $\exists \lambda \in \mathbb{R}$ mit $x = \lambda y$.

Gerade im Raum: $L^{(hom)} = \{ \begin{bmatrix} p1 \\ p2 \\ p3 \\ 1 \end{bmatrix} + \lambda \cdot \begin{bmatrix} v1 \\ v2 \\ v3 \\ 0 \end{bmatrix} | \lambda \in \mathbb{R}\}$

Für homogene Koordinaten in der BE gilt: $x \sim \Pi_0 \cdot L^{(hom)}$. $\Pi_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$

Definition 3.2. Das Urbild eines Punktes x in der BE ist die Menge der Punkte im Raum, die über die Projektion auf x abgebildet werden.
$preimage(x) = \Pi_0^{-1}(x)$

Feststellung: Urbild eines Punktes / einer Geraden sind Untervektorräume in $\mathbb{R}^3$.

Definition 3.3. Das Co-Bild eines Punktes / einer Geraden ist das orthogonale Komplement des Urbildes.

Bild, Urbild und Co-Bild von Punkten / Geraden sind einander eindeutig zuzuordnen, also äquivalente Darstellungen.

Bild = Urbild $\cap$ BE
Urbild = <Bild> = (Co-Bild)$^\perp$
Co-Bild = Urbild$^\perp$

Sei $\langle l \rangle$ das Co-Bild der Geraden L und sei $x \in \Pi(L)$, dann gilt: $x^T \cdot l = 0 = l^T \cdot x$.

Definition 3.4. Sei $l \in \mathbb{R}^3$, $l = [l_1 l_2 l_3]$, dann $\hat{l} \in \mathbb{R}^{3\times3}$, $\hat{l} = \begin{bmatrix} 0 & -l_3 & l_2 \\ l_3 & 0 & -l_1 \\ -l_2 & l_1 & 0 \end{bmatrix}$.

Es gilt: für $a \in \mathbb{R}^3 \quad \hat{l} \cdot a = l \times a$

$Ker\,\hat{l} = \langle l \rangle$
$Bild\,\hat{l} = Bild\,\hat{l}^T = (Ker\,\hat{l})^\perp = \langle l \rangle^\perp$

Es gilt:

	Bild	Urbild	Co-Bild
Punkt	$< x > \cap BE$	$< x >$	$< \hat{x} >$
Linie	$< \hat{l} > \cap BE$	$< \hat{l} >$	$< l >$

4 Epipolargeometrie

4.1 Euklidische Transformationen (Wdh.)

Rotationen: werden beschrieben durch Matrizen der speziellen orthogonalen Gruppe $SO(n) = \{X \in \mathbb{R}^{n \times n} | X^T X = I_n, detX = 1\}$

Bemerkung: ist $A = A^T$, so $\exists X \in SO(n)$ mit $X^T A X = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$

(weil $det[x_1, ..., -x_n] = -det[x_1, ..., x_n]$).
Bemerkung: $X_1, X_2 \in SO(n)$, dann:

1. $X_1 \cdot X_2 \in SO(n)$
2. (i.a.) $X_1 X_2 \neq X_2 X_1$
3. $X_1^T = X_1^{-1}$
4. $X_1^{-1} \in SO(n)$

Definition 4.1. Die Abbildung $g : \mathbb{R}^n \to \mathbb{R}^n, x \mapsto Rx + T$ mit $R \in SO(n), T \in \mathbb{R}^n$ heißt *euklidische Transformation.*

Bemerkung: $\|g(x) - g(y)\|_2 = \|x - y\|_2$ für alle $x, y \in \mathbb{R}^n$.

Beispiel 4.2. $g_1(x) = R_1 x + T_1, g_2 = R_2 x + T_2$
$g_2 \circ g_1(x) = R_2(R_1 x + T_1) + T_2 = \underbrace{R_2 R_1}_{\tilde{R}} x + \underbrace{R_2 T_1 + T_2}_{\tilde{T}}$ ist wieder eine euklidische Transformation.

In homogenen Koordinaten lässt sich g beschreiben durch Matrix-Vektor-Multiplikation, nämlich: $g(x^{(hom)}) = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} \cdot x^{(hom)}$

4.2 Epipolargeometrie (Kernstrahl-Geometrie)

Aufnahme eines Objektes aus verschiedenen Perspektiven.
-> Epipolargeometrie beschreibt die Beziehung zwischen den Bildern.

Annahme: Zwei identische kalibrierte Kameras mit relativer Pose (R, T): $g : x \mapsto Rx + T$
Wenn $X_1 \in \mathbb{R}^3$ Koordinaten eines Punktes p bezüglich Kameraframe 1,
und $X_2 \in \mathbb{R}^3$ Koordinaten desselben Punktes bezüglich Kameraframe 2,
dann gilt $X_2 = R \cdot X_1 + T$.

Ziel: Zusammenhang finden zwischen Bild von p in Kamera 1 und Kamera 2.

Seien x_1, x_2 die Bildpunkte von p in homogenen Koordinaten.
$\lambda_1 \cdot x_1 = X_1, \lambda_2 \cdot x_2 = X_2, \lambda_1, \lambda_2 \in \mathbb{R} \setminus \{0\}$, also:

$$\lambda_2 \cdot x_2 = R \cdot \lambda_1 \cdot x_1 + T \quad \Big| \times T (\to \hat{T}), T = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}, \hat{T} = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix}$$

$$\Rightarrow \lambda_2 \cdot \hat{T} \cdot x_2 = \hat{T} \cdot R \cdot \lambda_1 \cdot x_1 \quad \Big| \to x_2^T \qquad \big(x_2^T \cdot (T \times x_2) = 0\big)$$

$$\Rightarrow 0 = \lambda_1 x_2^T \hat{T} R x_1 \overset{\lambda_1 \neq 0}{\Rightarrow} \boxed{x_2^T \hat{T} R x_1 = 0}$$

Definition 4.3. Die Matrix $E := \hat{T} \cdot R \in \mathbb{R}^{3 \times 3}$ heißt *essentielle Matrix*. Die Gleichung $x_2^T \cdot \hat{T} R \cdot x_1 = 0$ heißt epipolare Bedingung (epipolar constraint).

Definition 4.4. Seien o_1, o_2 Ursprünge der beiden Koordinatensysteme.

1. Die Ebene, in der p, o_1, o_2 liegen, heißt *Epipolarebene* von p.
   (Falls $o_1 = 0$, so ist die Epipolarebene gegeben durch $\langle x1, T \rangle$).
2. $\Pi_1(o_2) =: e_1, \Pi_2(o_1) =: e_2$ heißen *Epipole*.
3. Der Schnitt der Epipolarebene von p mit der BE_1 (BE_2) ist eine Linie l_1 (l_2). Diese heißt *Epipolarlinie* von p.

Satz: Sei $E = \hat{T} \cdot R$ die essentielle Matrix, die die relative Pose zwischen den Kameras beschreibt.
x_1 Bildpunkt von p in Kamera 1, x_2 Bildpunkt von p in Kamera 2.
d.h. $x_2^T E x_1 = 0$. Dann gilt:

1. $e_2^T E = 0, E e_1 = 0$ (d.h. $e_2 \sim T, e_1 \sim R^T T$)
2. $l_2 \sim E x_1, l_1 \sim E^T x_2$
3. $l_i^T \cdot e_i = 0, l_i^T \cdot x_i = 0$ für $i = 1, 2$

4.3 Rechenbeispiele

Anmerkung: Verschiedene Koordinatensysteme beachten!

- T ist im Koordinatensystem 2.
- x_1, e_1, bzw. x_2, e_2 sind in ihrem jeweiligen Koordinatensystem (1 bzw. 2).
- $R \cdot x_1$, $R \cdot e_1$ bzw. $R^T \cdot x_2$, $R^T \cdot e_2$ sind die Punkte im jeweils anderen Koordinatensystem (2 bzw. 1).

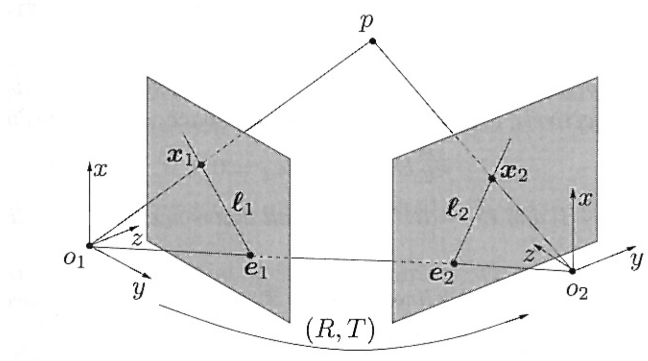

Abbildung 4.1: Epipolargeometrie

(1) Veranschaulichung der epipolaren Bedingung $\boxed{x_2^T \cdot \hat{T} \cdot R \cdot x_1 = 0}$

$x_2^T \cdot \hat{T} \cdot R \cdot x_1 = 0 \Leftrightarrow x_2 \perp T \times \underbrace{R \cdot x_1}_{\sim \overrightarrow{o_1 p} \text{ in KS 2}}$

Beobachtung: $dim \langle x_2, T, R \cdot x_1 \rangle = 2$ „...liegen in einer Ebene“. ✓(siehe Bild 4.1)

(2) Zeige: a) $Ee_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$ **und b)** $e_2^T E = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$

$E = \hat{T} \cdot R$
a) $E \cdot e_1 = \hat{T} \cdot R \cdot e_1 = T \times Re_1 = \vec{0}$ (T und Re_1 selbe Richtung, siehe Bild 4.1)

b) $(e_2^T E)^T = E^T e_2 = (R^T \hat{T}^T) e_2 = -R^T \hat{T} e_2 = -R^T \cdot (T \times e_2) = -R^T \vec{0} = \vec{0}$
$\Rightarrow e_2^T E = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}$

(3) Zeige: $l_1 \sim E^T x_2$ **und** $l_2 \sim Ex_1$

$l_1 = e_1 \times x_1$, also $l_1 \in \langle e_1, x_1 \rangle^\perp$
z.Z.: $E^T x_2 \in \langle e_1, x_1 \rangle^\perp$, also 1) $x_1 \perp E^T x_2$ und 2) $e_1 \perp E^T x_2$
1) $x_1^T E^T x_2 = x_2^T E x_1 \overset{\text{epip. Bed.}}{=} 0$ ✓
2) $e_1^T E^T x_2 = (Ee_1)^T x_2 \overset{\textbf{(2)a)}}{=} \vec{0}^T x_2 = 0$ ✓
$\Rightarrow E^T x_2 \in \langle e_1, x_1 \rangle^\perp \Rightarrow E^T x_2 \sim l_1$.

Für $l_2 \sim Ex_1$ analog.

(4) Zeige: $l_i^T e_i = 0$

Für $i = 1$:
$l_1 \sim E^T x_2$ **(3)** $\Leftrightarrow l_1 = \lambda \cdot (E^T x_2)$
$l_1^T e_1 = \lambda \cdot (E^T x_2)^T \cdot e_1 = \lambda x_2^T E e_1 \overset{\textbf{(2)a)}}{=} \lambda x_2^T \vec{0} = 0.$

Für $i = 2$ analog.

(5) Zeige: $l_i^T x_1 = 0$

Für $i = 1$:
$l_1^T x_1 = \lambda (E^T x_2)^T x_1 = \lambda x_2^T E x_1 \overset{\text{epip. Bed.}}{=} \lambda \cdot 0 = 0$

Für $i = 2$ analog.

5 8-Punkt-Algorithmus

Was können wir mit der essentiellen Matrix anfangen?

Drastische Reduktion des Suchraums für korrespondierende Punkte (Linie statt ganzes Bild).
Waagerechtes Ausrichten der Epipolarlinien.

5.1 Wiederholung: Mathe

Benötigte Werkzeuge:

- (Matrix-)Normen, konkret: Frobeniusnorm
- Kroneckerprodukt
- Singulärwertzerlegung

5.2 8-Punkt-Algorithmus

Gegeben $(x_i, y_i), i = 1...n$ korrespondierende Bildpunkte (in homogenen Koordinaten).
Aufgabe: Finde E.

Es gilt: $y_i^T \cdot E \cdot x_i \quad \forall i = 1...n.$ (*)
Bemerkung: E kann höchstens bis auf skalares Vielfaches bestimmt werden.

Umschreibe: $y^T E x = y^T \underbrace{[e_1, e_2, e_3]}_{E} x$

$$= y^T \cdot [t_1 e_1, t_2 e_2, t_3 e_3] \quad (\text{mit } x = \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix})$$

$$= [y^T \cdot t_1 \cdot e_1, y^T \cdot t_2 \cdot e_2, y^T \cdot t_3 \cdot e_3] = \underbrace{[y^T \cdot t_1, y^T \cdot t_2, y^T \cdot t_3]}_{\in \mathbb{R}^{1\times 9}} \underbrace{\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}}_{\in \mathbb{R}^9}$$

$$= (x \underbrace{\otimes}_{\text{Kroneckerprodukt}} y)^T \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}$$

(*) wird also zu:

$$D^T \cdot \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} = 0, \quad D = [x_1 \otimes y_1, ..., x_n \otimes y_2] \in \mathbb{R}^{9\times n}$$

(1) Wegen Skalierungsinvarianz verlange, dass $\|E\|_F = \left\| \begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix} \right\|_2 = 1.$

(2) Finde $\eta \in \mathbb{R}^q$ mit Norm $\|\eta\| = 1$ so, dass $\left\|D^T\eta\right\|_2$ minimal wird.

Lösung:
$min_{\|\eta\|=1} \left\|D^T\eta\right\|_2^2 = min_{\|\eta\|=1} \eta^T \cdot DD^T\eta \Rightarrow \eta$ ist normierter Eigenvektor von DD^T zum kleinsten Eigenwert.

nächstes Problem: projeziere die Matrix $H := [\eta_1, \eta_2, \eta_3]$ auf die näheste essentielle Matrix (d.h. auf eine mit Rang = 2 und 2 gleichen Singulärwerten).

Sei $H = U \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \sigma_3 \end{bmatrix} V^T$ SVD von H.

Erhalte Schätzung für E mittels:

$$E := U \begin{bmatrix} \hat{\sigma} & & \\ & \hat{\sigma} & \\ & & 0 \end{bmatrix} V^T \quad , \hat{\sigma} := \tfrac{1}{2}(\sigma_1 + \sigma_2)$$

bzw., wegen Skalierungsinvarianz:

$$E := U \begin{bmatrix} 1 & & \\ & 1 & \\ & & 0 \end{bmatrix} V^T$$

5.3 Lin. Alg. Review

Sei $A \in \mathbb{R}^{n\times k}, (k \leq n)$
Dann existiert $U \in O(n), V \in O(k)$ (orthogonale Matrizen) und $\sigma_1 \geq ... \geq \sigma_n \geq 0$ mit

$$A = U \cdot \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \\ & 0 & \end{bmatrix} V^T \text{ (Singulärwertzerlegung SVD)}$$

Bemerkung: $V^T(A^TA)V = \begin{bmatrix} \sigma_1^2 & & 0 \\ & \ddots & \\ 0 & & \sigma_n^2 \end{bmatrix}, U^T(AA^T)U = \begin{bmatrix} \sigma_1^2 & & & & 0 \\ & \ddots & & & \\ & & \sigma_n^2 & & \\ & & & 0 & \\ & & & & \ddots \\ 0 & & & & 0 \end{bmatrix}$

* i.a. nur numerische Approximation bestimmbar.

* U, V nicht eindeutig, aber $\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix}$ schon!

Satz: $E \in \mathbb{R}^{3\times 3} \setminus \{0\}$ ist essentielle Matrix genau dann, wenn

$E = U \cdot \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} \cdot V^T$, mit $U, V \in SO(3)$. (Beweis: Ma [1], Theorem 5.5)

Beispiel: SVD von $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} =: A$

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \mp 1 \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \pm 1 \end{bmatrix}}_{V^T}$$

Achtung: Entweder U oder $V \notin SO(3)$ ($det = -1$).
Aber: $U, V \in SO(3)$ immer konstruierbar, z.B. für $U \notin SO(3)$:

$E = U \cdot \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} \cdot V^T = U \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix} V^T$. (analog für V)

d.h. wir können annehmen, dass $U, V \in SO(3)$.
* $\mathbb{R}^{n\times k}$ ist reeller Vektorraum mit Skalarprodukt.
$\mathbb{R}^{n\times k} \times \mathbb{R}^{n\times k} \longmapsto \mathbb{R}, (A, B) \mapsto tr(A \times B^T)$

* $tr(A \times B^T) = tr(B^T \times A) = tr(B \times A^T) = tr(A^T \times B)$
dadurch ist Norm erklärt: $\|A\|_F = [tr(A^T A)]^{1/2}$ (Frobenius-Norm)

* $\left\|UAV^T\right\|_F = \|A\|_F$ für $U \in O(n), V \in O(n)$
$\Rightarrow \|A\|_F = (\Sigma|\sigma_i^2)^{1/2}, \sigma_i$ Singulärwerte von A.

Satz: Sei $A \in \mathbb{R}^{n\times k}$ mit SVD $A = U \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T, \Sigma = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix}$

Sei $M^{(l)} = \{X \in \mathbb{R}^{n\times k} | rk(X) \leq l\}$

Sei $A^{trunc} := U \cdot \begin{bmatrix} \sigma_1 & & & & & \\ & \ddots & & & & \\ & & \sigma_l & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \\ & & 0 & & & \end{bmatrix} V^T$

Dann gilt:
1) $A^{trunc} \in M^{(l)}$
2) Für alle $X \in M^{(l)}$ gilt: $\|A - X\|_F \geq \left\|A - A^{trunc}\right\|_F$

5.4 Pose Recovery

Wir haben E, wie bekommen wir nun $R, \hat{T}$?

Sei $E = U \cdot \underbrace{\begin{bmatrix} \sigma & & \\ & \sigma & \\ & & 0 \end{bmatrix}}_{\Sigma} \cdot V^T, \quad U, V \in SO(3).$

Sei $\hat{T}_1 = U \cdot \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Sigma \cdot U^T$, $R_1 = U \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot V^T$
und $\hat{T}_2 = U \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Sigma \cdot U^T$, $R_2 = U \cdot \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot V^T$
Es gibt genau zwei euklidische Transformationen die zu E gehören, nämlich:
$(\hat{T}_1, R_1), (\hat{T}_2, R_2)$. (Beweis in Ma [1])

Bemerkung: Erhält man also E über 8-Punkt-Algorithmus, ist $\hat{T}$ bis auf Skalierung bestimmt.
Festlegen von $\left\|\hat{T}\right\|_F = 1 \quad \hat{=}$ Festlegen der Längeneinheit.

Gegeben (durch 8-Punkt-Algo) also $\pm E$. $\quad E = U\Sigma V^T$
SVD von $-E = U(-\Sigma)V^T = \underbrace{U \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}}_{\tilde{U}} \Sigma V^T$

$\hat{T}_3 := \tilde{U} \cdot \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Sigma \cdot \tilde{U}^T = \hat{T}_1$
$R_3 = \tilde{U} \cdot \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \Sigma \cdot V^T = R_2$
$\hat{T}_4 = \hat{T}_2, R_4 = R_1$

also 4 euklidische Bewegungen:
$(\hat{T}_1, R_1), (\hat{T}_2, R_2) \qquad (\hat{=} E)$
$(\hat{T}_1, R_2), (\hat{T}_2, R_1) \qquad (\hat{=} -E)$

Nur ein Paar erfüllt die Gleichung $\lambda_2 y = \lambda_1 R x + T$ mit $\lambda_1, \lambda_2 > 0$.
$\rightarrow$ Ausprobieren.

Literaturverzeichnis

[1] Yi Ma, Stefano Soatto, Jana Kosecka, and S. Shankar Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer Verlag.